目录

准备工作

在您开始使用日志服务进行数据采集、查询分析等操作时，需先完成本文中的以下几个准备工作。

开通KLog服务

新用户在使用KLog前，请先进入金山云KLog控制台自行开通服务。开通服务，代表同意日志服务KLog的服务协议。

创建项目

1. 登录金山云日志服务Klog控制台。
2. 点击导航**项目列表**，进入到**项目列表页**。
3. 点击**创建项目**，填写项目相关信息。
   - 项目名称：参照页面中的名称格式要求，定义项目名称，名称保存后不支持修改。以”project_demo“为例。
   - 地域：目前KLog仅支持**华北1（北京）**。

创建日志池

1. 登录金山云日志服务Klog控制台。
2. 点击导航**项目列表**，进入到项目列表页。
3. 在项目列表页中，点击项目名称，进入到**项目>概览页**。
4. 点击**创建日志池**，填写日志池相关信息。
   - 名称：参照页面中的名称格式要求，定义日志池名称，名称保存后不支持修改。以”logpool_demo“为例。
   - 保存时间：指日志数据存储时长，KLog会自动删除超出存储时长的数据，时长范围支持1-3000天
   - 分区数：分区数越多，代表数据读写吞吐量越大，请根据业务数据情况合理设置分区数。

数据采集

本文通过介绍如何使用Filebeat采集云主机KEC的Nginx日志数据为例，说明如何实现数据实时采集。当前除了兼容开源Filebeat采集外，还支持API方式实时推送数据。详见API上传数据。

前提条件

在华北1（北京）已创建云服务器实例，且服务器上有待采集的日志数据。

安装Filebeat采集器

1. 下载KLog提供的Filebeat采集器。
   - 最新Linux版本的klog-filebeat可以点击此处下载。
   - 或直接在您的服务器上下载：

     ```
     wget "https://ks3-cn-beijing.ksyun.com/klog/filebeat/klog-filebeat.latest.tar.gz"
     ```

2. 安装Filebeat采集器，把安装包解压缩到您需要的路径即可。

   ```
   tar xvf klog-filebeat.latest.tar.gz
   ```

配置采集任务

接下来以采集/var/log/nginx/access.log日志文件为例，配置数据采集任务。 日志示例如下：

```
10.168.255.134 [01/Sep/2016:11:04:48 +0800] "GET /account/fund/fundDetail.html?1472699086917 HTTP/1.1" - 200 3777 "https://wenjinbao.winfae.com/account/myAccount.html" "Mozilla/5.0 (Windows NT 5.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/47.0.2526.108 Safari/537.36 2345Explorer/7.1.0.12633" 0.000 115.226.250.21
```

字段格式定义如下：

```
log_format  main  '$remote_addr [$time_local] "$request" '
                    '$request_body $status $body_bytes_sent "$http_referer" "$http_user_agent" '
                    '$request_time $http_x_forwarded_for';
```

klog-filebeat需要2个配置文件，分别是filebeat.yml和credential.ini，这两个配置文件均在klog-filebeat压缩包中，解压后即可看到。

credential.ini

这个配置文件的内容是您的金山云access_key和secret_key。内容示例：

```
[klog]
access_key = AKLT6-bcdesfg-ajsIjfiejI9
secret_key = Jf2390j9E9finfiDIOFJFD8483+dfsDVdOCTFazCnDEAw+mxA/7Lfeh3ugErwoKKb5wNOIei==
```

filebeat.yml

这个是主配置文件，内容示例如下：

```
filebeat.inputs:
- type: log
  enabled: true
  paths:
    # 需要采集的日志文件
    - /var/log/nginx/access.log

output.klog:
  endpoint: https://klog-cn-beijing.ksyun.com
  credential:
    path: credential.ini
    check_interval: 60
  targets:
    - project_name: project_demo
      pool_name: project_demo
      fields:
        - key: message
```

```
          skip_root: true
        - key: grok
          skip_root: true

  processors:
  #grok处理器可以把普通文本日志解析为json对象。
  #由于grok是基于正则表达式的，所以开启该选项会增加filebeat的资源消耗。
  - grok:
      # 解析哪个字段
      source_field: message

      # 解析结果保存到哪个字段
      target_field: grok

      # 用于解析的表达式
      pattern: '%{IPORHOST:clientip} \[%{HTTPDATE:time}\] \"%{WORD:verb} %{URIPATHPARAM:request} HTTP/%{NUMBER:httpversion}\" \- %{NUMBER:http_status_code} %{NUMBER:
bytes} \"(?<http_referer>\S+)\" \"(?<http_user_agent>(\S+\s+)*\S+)\" (%{BASE16FLOAT:request_time}) (%{IPORHOST:http_x_forwarded_for}|-)"
```

关于output.klog的详细配置说明可参考klog-filebeat。其余详细配置可参考filebeat官方文档。

注意：在输出到klog的情况下，filebeat的modules功能无法使用。

启动采集器

执行如下命令，启动klog-filebeat。启动后，会在指定的日志池中看到写入的日志数据。

```
./filebeat -e
```

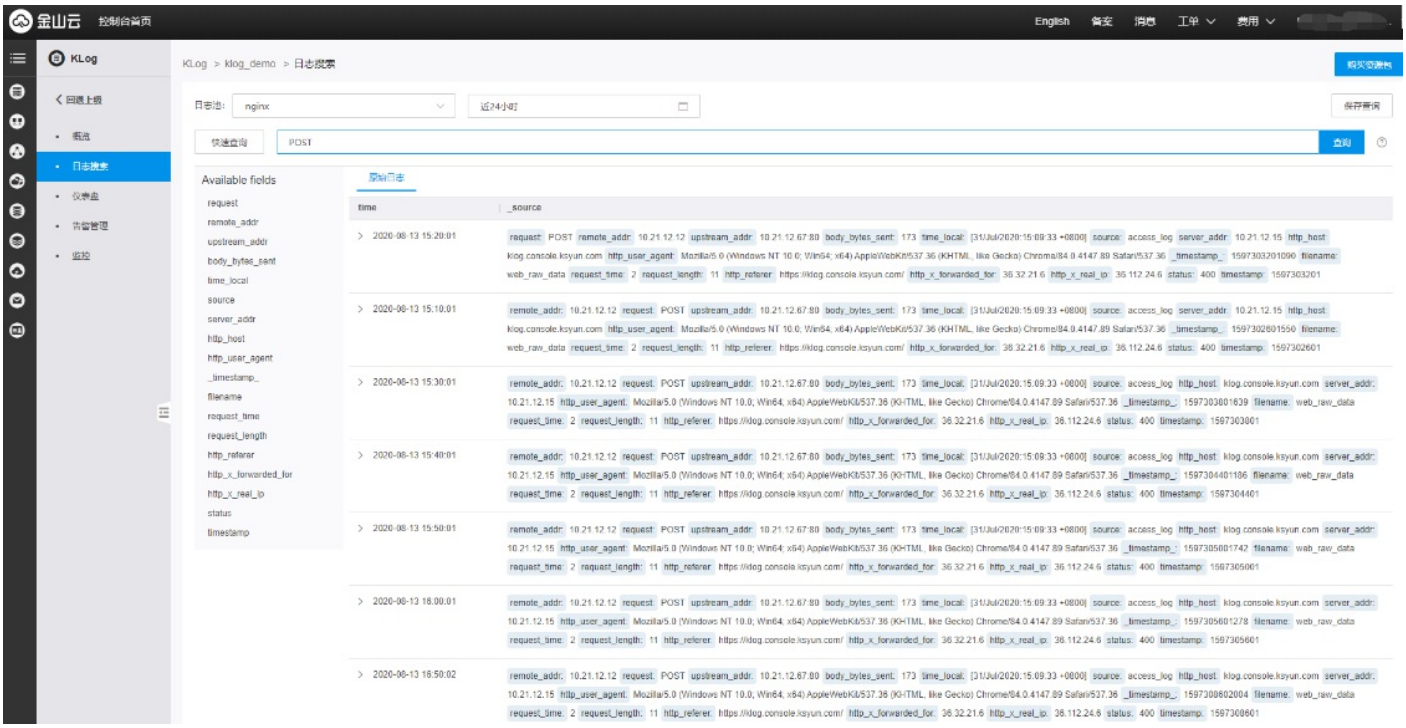
日志搜索与可视化

日志搜索

进入KLog控制台，点击左侧导航中的**日志搜索**或者点击**日志池列表**中的**搜索**，进入到日志搜索页面。 日志服务支持全文检索、键值检索、SQL查询等查询语法，接下来结合示例日志举例介绍这几种日志搜索。

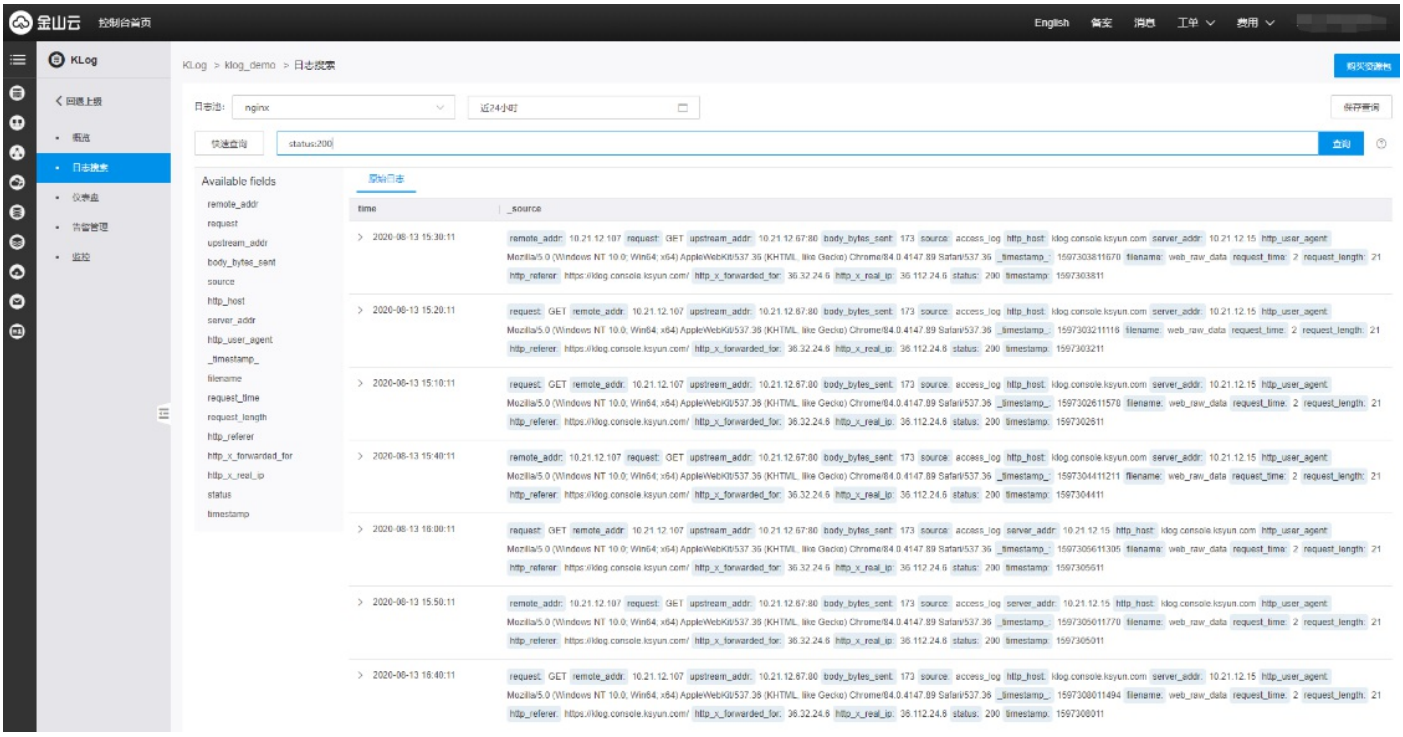
全文检索

示例：查询近24小时，日志中包含POST的日志数据。 如下图，日志池选中”nginx“，时间范围选中”近24小时“，输入框中输入”POST“（区分大小写），点击查询，查询结果如下图。



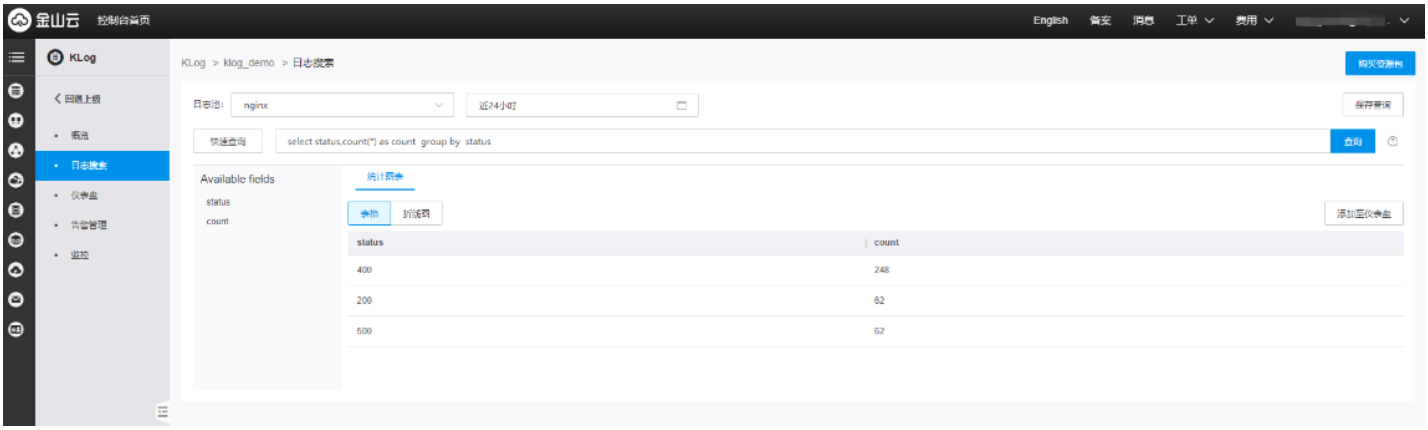
键值查询

示例：查询近24小时，status为200的日志数据。 如下图，日志池选中”nginx“，时间范围选中”近24小时“，输入框中输入”status:200“,点击查询，查询结果如下图。



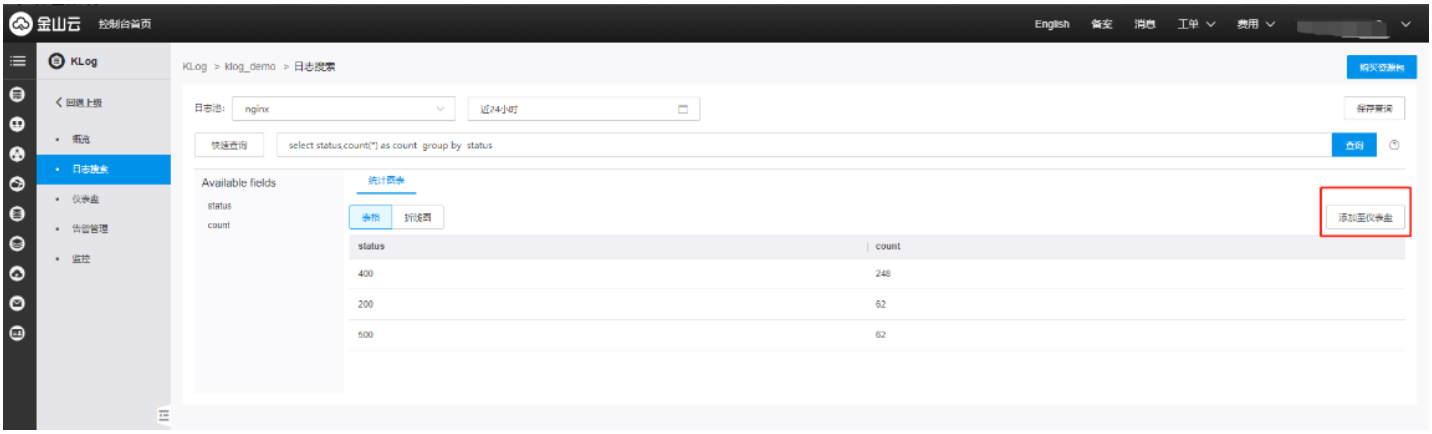
SQL查询

示例：统计近24小时status各个值的日志条数。 如下图，日志池选中”nginx“，时间范围选中”近24小时“，输入框中输入”select status,count(*) as count group by status“,点击查询，查询结果如下图。



添加图表到仪表盘

点击统计图表右侧的”添加至仪表盘“。



如下图，填写仪表盘名称，或者选择已有仪表盘，把上面的表格添加到仪表盘中。


保存至仪表盘
操作类型：
新建仪表盘
选择已有仪表盘
仪表盘名称：
图表名称：